

0570  
0507

#7



OIE

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/020,674

DATE: 05/02/2002

TIME: 15:37:07

Input Set : A:\06501-092001.TXT

Output Set: N:\CRF3\05022002\J020674.raw

4 <110> APPLICANT: Yamamoto, Hiroaki  
 5 Onodera, Keiko  
 6 Tani, Yoshiki  
 8 <120> TITLE OF INVENTION: NOVEL (R)-2,3-BUTANEDIOL DEHYDROGENASE  
 11 <130> FILE REFERENCE: 06501-092001  
 13 <140> CURRENT APPLICATION NUMBER: 10/020,674  
 C--> 14 <141> CURRENT FILING DATE: 2002-04-19  
 16 <150> PRIOR APPLICATION NUMBER: JP 2000-333363  
 17 <151> PRIOR FILING DATE: 2000-10-31  
 19 <160> NUMBER OF SEQ ID NOS: 17  
 21 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 1143  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Pichia angusta  
 28 <400> SEQUENCE: 1  
 29 atgaaagggtt tactttatta cggtagaaac gatattcgt actccgaaac ggttcctgaa 60  
 30 ccggagatca agaatcccaa cgaatgtcaag atcaaagtca gctattgtgg aatctgtggc 120  
 31 acggacttga aagaattcac atattctgga ggtcctgttt tttccctaa acaaggcacc 180  
 32 aaggacaaga tttcgggata cgaacttcct ctctgtcctg gacatgaatt tagcggaaacg 240  
 33 gtggtcgagg ttggtctctg tgtcacaagt gtgaaacctg gtgacagagt cgagttgaa 300  
 34 gctacgtcgc attgtctcga cagatcgccg tacaaggaca cggtcgcccc agaccttggg 360  
 35 ctctgtatgg cctgccagag cggatctccg aactgctgtg cgtcgtctgag cttctgcggt 420  
 36 ttgggtggtg ccagcggcgg ttttgccgag tacgtcgttt acggtgagga ccacatggtc 480  
 37 aagctgccag actcgattcc cgaagatatt ggagcactgg ttgagcctat ttctgttgcc 540  
 38 tggcatgctg ttgaacgcgc tagattccag cctggtcaga cggccctggt tcttgaggga 600  
 39 ggtcctatcg gccttgccac cattcttgct ctgcaaggcc atcatgcggg caaaattgtg 660  
 40 tgttccgagc cggccttgat cagaagacag tttgcaaagg aactgggcgc tgaagtgttc 720  
 41 gatccttcta catgtgacga cgcaaagtgt gttctcaagg ctatggtgcc ggagaacgag 780  
 42 ggattccatg cagccttcga ctgctctggt gttcctcaga cattcaccac ctcaattgtc 840  
 43 gccacgggac cttctggaat cgccgtcaat gtggccgttt ggggagacca cccaattgga 900  
 44 ttcatgccaa tgtctctgac ttaccaggag aaatacgcta ccggtcccat gtgctacacc 960  
 45 gtcaaggact tccaggaagt tgtcaaggcc ttggaagatg gtctcatatc tttggacaaa 1020  
 46 gcgcgcaaga tgattacagg caaagtcac ctaaaggacg gagtcgagaa gggctttaa 1080  
 47 cagctgatcg agcacaagga gaacaatgtc aagatcctgg tgacgccgaa cgaggtttcc 1140  
 48 taa 1143  
 50 <210> SEQ ID NO: 2  
 51 <211> LENGTH: 380  
 52 <212> TYPE: PRT  
 53 <213> ORGANISM: Pichia angusta  
 55 <400> SEQUENCE: 2  
 56 Met Lys Gly Leu Tyr Tyr Gly Thr Asn Asp Ile Arg Tyr Ser Glu  
 57 1 5 10 15

## RAW SEQUENCE LISTING

DATE: 05/02/2002

PATENT APPLICATION: US/10/020,674

TIME: 15:37:07

Input Set : A:\06501-092001.TXT

Output Set: N:\CRF3\05022002\J020674.raw

```

58 Thr Val Pro Glu Pro Glu Ile Lys Asn Pro Asn Asp Val Lys Ile Lys
59                20                25                30
60 Val Ser Tyr Cys Gly Ile Cys Gly Thr Asp Leu Lys Glu Phe Thr Tyr
61                35                40                45
62 Ser Gly Gly Pro Val Phe Phe Pro Lys Gln Gly Thr Lys Asp Lys Ile
63                50                55                60
64 Ser Gly Tyr Glu Leu Pro Leu Cys Pro Gly His Glu Phe Ser Gly Thr
65 65                70                75                80
66 Val Val Glu Val Gly Ser Gly Val Thr Ser Val Lys Pro Gly Asp Arg
67                85                90                95
68 Val Ala Val Glu Ala Thr Ser His Cys Ser Asp Arg Ser Arg Tyr Lys
69                100               105               110
70 Asp Thr Val Ala Gln Asp Leu Gly Leu Cys Met Ala Cys Gln Ser Gly
71                115               120               125
72 Ser Pro Asn Cys Cys Ala Ser Leu Ser Phe Cys Gly Leu Gly Gly Ala
73                130               135               140
74 Ser Gly Gly Phe Ala Glu Tyr Val Val Tyr Gly Glu Asp His Met Val
75 145               150               155               160
76 Lys Leu Pro Asp Ser Ile Pro Asp Asp Ile Gly Ala Leu Val Glu Pro
77                165               170               175
78 Ile Ser Val Ala Trp His Ala Val Glu Arg Ala Arg Phe Gln Pro Gly
79                180               185               190
80 Gln Thr Ala Leu Val Leu Gly Gly Gly Pro Ile Gly Leu Ala Thr Ile
81                195               200               205
82 Leu Ala Leu Gln Gly His His Ala Gly Lys Ile Val Cys Ser Glu Pro
83                210               215               220
84 Ala Leu Ile Arg Arg Gln Phe Ala Lys Glu Leu Gly Ala Glu Val Phe
85 225               230               235               240
86 Asp Pro Ser Thr Cys Asp Asp Ala Asn Ala Val Leu Lys Ala Met Val
87                245               250               255
88 Pro Glu Asn Glu Gly Phe His Ala Ala Phe Asp Cys Ser Gly Val Pro
89                260               265               270
90 Gln Thr Phe Thr Thr Ser Ile Val Ala Thr Gly Pro Ser Gly Ile Ala
91                275               280               285
92 Val Asn Val Ala Val Trp Gly Asp His Pro Ile Gly Phe Met Pro Met
93                290               295               300
94 Ser Leu Thr Tyr Gln Glu Lys Tyr Ala Thr Gly Ser Met Cys Tyr Thr
95 305               310               315               320
96 Val Lys Asp Phe Gln Glu Val Val Lys Ala Leu Glu Asp Gly Leu Ile
97                325               330               335
98 Ser Leu Asp Lys Ala Arg Lys Met Ile Thr Gly Lys Val His Leu Lys
99                340               345               350
100 Asp Gly Val Glu Lys Gly Phe Lys Gln Leu Ile Glu His Lys Glu Asn
101                355               360               365
102 Asn Val Lys Ile Leu Val Thr Pro Asn Glu Val Ser
103                370               375               380
105 <210> SEQ ID NO: 3
106 <211> LENGTH: 10
107 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/020,674

DATE: 05/02/2002

TIME: 15:37:07

Input Set : A:\06501-092001.TXT

Output Set: N:\CRF3\05022002\J020674.raw

```

108 <213> ORGANISM: Pichia angusta
110 <400> SEQUENCE: 3
111 Lys Pro Gly Asp Arg Val Ala Val Glu Ala
112 1 5 10
114 <210> SEQ ID NO: 4
115 <211> LENGTH: 21
116 <212> TYPE: PRT
117 <213> ORGANISM: Pichia angusta
119 <400> SEQUENCE: 4
120 Ala Thr Ser His Cys Ser Asp Arg Ser Arg Tyr Lys Asp Thr Val Ala
121 1 5 10 15
122 Gln Asp Leu Gly Leu
123 20
125 <210> SEQ ID NO: 5
126 <211> LENGTH: 6
127 <212> TYPE: PRT
128 <213> ORGANISM: Pichia angusta
130 <400> SEQUENCE: 5
131 Phe His Ala Ala Phe Asp
132 1 5
134 <210> SEQ ID NO: 6
135 <211> LENGTH: 20
136 <212> TYPE: DNA
137 <213> ORGANISM: Artificial Sequence
139 <220> FEATURE:
140 <223> OTHER INFORMATION: Artificially synthesized primer sequence
W--> 142 <221> NAME/KEY: misc_feature
143 <222> LOCATION: 6, 9, 15, 18
144 <223> OTHER INFORMATION: n = a, c, g, or t
W--> 146 <400> 6
W--> 147 aarccnggng aymgngtngc 20
149 <210> SEQ ID NO: 7
150 <211> LENGTH: 20
151 <212> TYPE: DNA
152 <213> ORGANISM: Artificial Sequence
154 <220> FEATURE:
155 <223> OTHER INFORMATION: Artificially synthesized primer sequence
W--> 157 <221> NAME/KEY: misc_feature
158 <222> LOCATION: 9, 12
159 <223> OTHER INFORMATION: n = a, c, g, or t
W--> 161 <400> 7
W--> 162 tctrcraang cngcrtgraa 20
164 <210> SEQ ID NO: 8
165 <211> LENGTH: 530
166 <212> TYPE: DNA
167 <213> ORGANISM: Pichia angusta
169 <400> SEQUENCE: 8
170 aagccgggtg atcgtgtcgc agttgaagct acgtcgcatt gctccgacag atcgcgctac 60
171 aaggacacgg tcgcccaaga cttgggctc tgtatggcct gccagagcgg atctccgaac 120

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/020,674

DATE: 05/02/2002

TIME: 15:37:07

Input Set : A:\06501-092001.TXT

Output Set: N:\CRF3\05022002\J020674.raw

```

172 tgctgtgcgt cgctgagctt ctgcggtttg ggtggtgccg gcggcggttt tgccgagtac 180
173 gtogttttacg gtgaggacca catggtcaag ctgccagact cgattcccga cgatattgga 240
174 gcaactggttg agcctatttc tgttgccctg catgctgttg aacgcgctag attccagcct 300
175 ggtcagacgg ccctggttct tggaggaggt cctatcgccg ttgccaccat tcttgctctg 360
176 caaggccatc atgcgggcaa aattgtgtgt tccgagccgg ccttgatcag aagacagttt 420
177 gcaaaggaac tgggcgctga agtgttcgat ccttctacat gtgacgacgc aaatgctgtt 480
178 ctcaaggcta tggtgccgga gaacgaggga ttccacgccg ccttcgatga 530
180 <210> SEQ ID NO: 9
181 <211> LENGTH: 26
182 <212> TYPE: DNA
183 <213> ORGANISM: Artificial Sequence
185 <220> FEATURE:
186 <223> OTHER INFORMATION: Artificially synthesized primer sequence
190 <400> SEQUENCE: 9
191 ttggcatgcg atctgtcggg gcaatg 26
193 <210> SEQ ID NO: 10
194 <211> LENGTH: 27
195 <212> TYPE: DNA
196 <213> ORGANISM: Artificial Sequence
198 <220> FEATURE:
199 <223> OTHER INFORMATION: Artificially synthesized primer sequence
203 <400> SEQUENCE: 10
204 tgagcatgca aatgctgttc tcaaggc 27
206 <210> SEQ ID NO: 11
207 <211> LENGTH: 107
208 <212> TYPE: DNA
209 <213> ORGANISM: Pichia angusta
211 <400> SEQUENCE: 11
212 gaatttagcg gaacgggtgt cgagggttggc tctggtgtca caagtgtgaa acctggtgac 60
213 agagtgcgag ttgaagctac gtgcgattgc tccgacagat cgcattgc 107
215 <210> SEQ ID NO: 12
216 <211> LENGTH: 706
217 <212> TYPE: DNA
218 <213> ORGANISM: Pichia angusta
220 <400> SEQUENCE: 12
221 gcatgcaaat gctgttctca aggctatggt gccggagaac gagggattcc atgcagcctt 60
222 cgactgctct ggtgttcttc agacattcac cacctcaatt gtgccacgga gaccttctgg 120
223 aatcgccgct aatgtggcgg tttggggaga ccaccaatt ggattcatgc caatgtctct 180
224 gacttaccag gagaaatacg ctaccggctc catgtgtctac accgtcaagg acttccagga 240
225 agttgtcaag gccttgaag atggtctcat atctttggac aaagcgcgca agatgattac 300
226 aggcaaagtc cacctaaagg acggagtcca gaagggcttt aaacagctga tcgagcacia 360
227 ggagaacaat gtcaagatcc tggtgacgcc gaacgaggtt tctaactaa taatatacat 420
228 acatcataca tatgtatgtc cttagagcaa gacttgcgca ttaggaaaaa tagctggtag 480
229 tttgcattat ggtggcgggc ctcccaggaa attaacttat gatttacata tggactcgat 540
230 tacgtaacag gtgctgagca ttttaataatt acctactatt ttctaaatta gtaaattgta 600
231 tgtttcttga gcaggaggag atactagagc aatttcaaaa catctccaat tgccaaatcc 660
232 ctgtgtccga acagattgca ttgctagagt ctgtgaactg gaattt 706
234 <210> SEQ ID NO: 13
235 <211> LENGTH: 620

```

## RAW SEQUENCE LISTING

DATE: 05/02/2002

PATENT APPLICATION: US/10/020,674

TIME: 15:37:07

Input Set : A:\06501-092001.TXT

Output Set: N:\CRF3\05022002\J020674.raw

```

236 <212> TYPE: DNA
237 <213> ORGANISM: Pichia angusta
239 <400> SEQUENCE: 13
240 tgacattcca caccaacttc tgccgccacc actgcaatcc tgtaggcgaa caggacgatg      60
241 caggactatt tctctatttt ttcccatcgt gcaccctgaa ccaatacggg ggaggcatgg      120
242 gaattttccg cgctaattcca gtcaacggta acaagaccag gatggagtgt gaatatttct      180
243 ttgacggcag cgatgaggag ttcgaggcct acttcaagtt tgccagacag gtcgcaactcg      240
244 aggatatttg gctgtgtgag gcggcccaac agaaccttat aagtggggtg taccaacagg      300
245 gcttgctgca tcctaaaaaa gaagtcgggg tggtttacta ccagtcgctg gttcgtgaaa      360
246 gaataatggc ttagctccga gatgtggagg cagtctggtc agactgtgcg gcaattaaat      420
247 aagacgcgga tgtactgcac cagagtgaat aaaggaattc caattcgata gcaaataattg      480
248 ctgtaataat gagtgaccag atttattacc gaacctagcc agcccggggt tttttacaca      540
249 ataggaaaaa aaggactcga ttattcgatg ctgctgcaaa tcacgccaga cataataagt      600
250 caccggttta ctccgcatgc
252 <210> SEQ ID NO: 14
253 <211> LENGTH: 30
254 <212> TYPE: DNA
255 <213> ORGANISM: Artificial Sequence
257 <220> FEATURE:
258 <223> OTHER INFORMATION: Artificially synthesized primer sequence
260 <400> SEQUENCE: 14
261 tgacctgcagc gccagacata ataagtcacc      30
263 <210> SEQ ID NO: 15
264 <211> LENGTH: 523
265 <212> TYPE: DNA
266 <213> ORGANISM: Pichia angusta
268 <400> SEQUENCE: 15
269 ctgcagcgcc agacataata agtcaccogt ttactccgca tgcaactccc cactgatcat      60
270 gattaatggt tctggacggc taaatcattg atcactgcgt cccggacctc gtaccgacgt      120
271 ggaaattagc cggcactcgg ttgtgagaga ttatctata taaaccacaa aatcctatct      180
272 cccttttgcc aatgaaagggt ttactttatt acggtacaaa cgatattcgc tactccgaaa      240
273 cggttcctga accggagatc aagaatccca acgatgtcaa gatcaaagtc agctattgtg      300
274 gaatctgtgg cacggacttg aaagaattca catattctgg aggtcctggt tttttcccta      360
275 aacaaggcac caaggacaag atttcgggat acgaacttcc tctctgtcct ggacatgaat      420
276 ttagcggaac ggtggtcgag gttggctctg gtgtcacaag tgtgaaacct ggtgacagag      480
277 tcgcagttga agctacgtcg cattgtctcg acagatcgca tgc
279 <210> SEQ ID NO: 16
280 <211> LENGTH: 30
281 <212> TYPE: DNA
282 <213> ORGANISM: Artificial Sequence
284 <220> FEATURE:
285 <223> OTHER INFORMATION: Artificially synthesized primer sequence
287 <400> SEQUENCE: 16
288 tgctcatgaa aggtttactt tattacggta      30
290 <210> SEQ ID NO: 17
291 <211> LENGTH: 28
292 <212> TYPE: DNA
293 <213> ORGANISM: Artificial Sequence
295 <220> FEATURE:

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 05/02/2002  
PATENT APPLICATION: US/10/020,674      TIME: 15:37:08

Input Set : A:\06501-092001.TXT  
Output Set: N:\CRF3\05022002\J020674.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:6; N Pos. 6,9,15,18

Seq#:7; N Pos. 9,12

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/020,674

DATE: 05/02/2002

TIME: 15:37:08

Input Set : A:\06501-092001.TXT

Output Set: N:\CRF3\05022002\J020674.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:142 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:146 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6  
L:147 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0  
L:157 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:161 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:7  
L:162 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0